MOLECULAR ECOLOGY

Molecular Ecology (2012) **21**, 5138–5150 doi: 10.1111/j.1365-294X.2012.05759.x

Deep sequencing reveals extensive variation in the gut microbiota of wild mosquitoes from Kenya

J. OSEI-POKU,* C. M. MBOGO,† W. J. PALMER* and F. M. JIGGINS*

Department of Genetics, University of Cambridge, Downing Street, Cambridge, CB2 3EH, UK, †Kenya Medical Research Institute (KEMRI), Centre for Geographic Medicine Research, Coast, P.O. Box 428, Kilifi 80108, Kenya

Abstract

The mosquito midgut is a hostile environment that vector-borne parasites must survive to be transmitted. Commensal bacteria in the midgut can reduce the ability of mosquitoes to transmit disease, either by having direct anti-parasite effects or by stimulating basal immune responses of the insect host. As different bacteria have different effects on parasite development, the composition of the bacterial community in the mosquito gut is likely to affect the probability of disease transmission. We investigated the diversity of mosquito gut bacteria in the field using 454 pyrosequencing of 16S rRNA to build up a comprehensive picture of the diversity of gut bacteria in eight mosquito species in this population. We found that mosquito gut typically has a very simple gut microbiota that is dominated by a single bacterial taxon. Although different mosquito species share remarkably similar gut bacteria, individuals in a population are extremely variable and can have little overlap in the bacterial taxa present in their guts. This may be an important factor in causing differences in disease transmission rates within mosquito populations.

Keywords: *Aeromonas*, *Asaia*, gut bacteria, natural populations of mosquitoes, pyrosequencing

Received 25 April 2012; revision received 4 July 2012; accepted 15 July 2012

Introduction

To be transmitted, mosquito-borne parasites must penetrate the insect midgut before completing their development in the tissues of the insect. However, the midgut is a hostile environment in which many parasites perish; for example, in *Anopheles* mosquitoes, only a small minority of *Plasmodium* parasites survive the midgut (Al-Olayan *et al.* 2002). Therefore, understanding the factors that are affecting the survival of parasites in the mosquito midgut has the potential to allow us to reduce or even block disease transmission. One important factor is the innate immune system of the insect, which can be activated when parasites such as *Plasmodium* and filarial parasites invade the midgut (Osta *et al.* 2004; Michel & Kafatos 2005; Erickson *et al.* 2009; Michalski *et al.* 2010), resulting in the upregulation of immune genes such as those encoding antimicrobial peptides and thioester-containing proteins that have anti-parasite effects (Blandin *et al.* 2004; Richman *et al.* 1997; Vlachou *et al.* 2005). However, it has recently become clear that bacteria living in the insect gut can also have an important role.

The mosquito gut is naturally inhabited by a community of bacteria that can disrupt the development of human parasites such as *Plasmodium* (Pumpuni *et al.* 1993, 1996; Straif *et al.* 1998; Gonzalez-Ceron *et al.* 2003; Dong *et al.* 2009). One way in which gut bacteria can interfere with parasite development is exerting direct anti-parasite effects (reviewed in Azambuja *et al.* 2005). This has been shown in the case of an *Enterobacter* bacterium isolated from African populations of *Anopheles arabiensis*, which generates reactive oxygen species that make the mosquitoes resistant to *Plasmodium* infection (Cirimotich *et al.* 2011). The presence of gut microbiota has also been shown to activate the immune response of mosquitoes, causing the release of immune proteins that are cross-reactive with the parasites (Dong *et al.* 2009), and this may indirectly block the development of parasites such as *Plasmodium*.

Correspondence: Jewelna Osei-Poku, Fax: +44 1223 333992 ; E-mail: j.osei-poku@gen.cam.ac.uk

The different bacteria that have been isolated from mosquito guts can have dramatically different effects on the development of human parasites. For example, Cirimotich *et al.* (2011) isolated four bacterial species from wild *An. arabiensis* mosquitoes and found that when these were fed to mosquitoes in the laboratory, an *Enterobacter* sp. almost completely inhibited *Plasmodium* development, while the bacterium *Bacillus pumilus* had no effect. As these are natural gut bacteria, it is therefore possible that the composition of the gut microbiota might have an important impact on rates of disease transmission in the wild and cause differences in the rate that different species or populations of vectors transmit disease. Furthermore, if the composition of the gut microbiota could be manipulated, then this could be a method of disease control. For example, sugar-bait methods that are used for killing mosquitoes (Müller *et al.* 2010) could be adapted for infecting adult mosquitoes with specific bacteria.

To predict how gut bacteria will affect disease transmission, it is important to investigate how the community of gut bacteria varies across different mosquito species, populations and individuals. One approach is to culture the bacteria isolated from the gut and characterize the different isolates. It is common, however, to find that the majority of bacteria in environmental samples cannot be cultured, so this approach may give a false representation of the bacterial species present and their relative abundance. A less biased approach is to amplify the bacterial 16S rRNA gene by PCR and then clone and sequence the PCR product. This approach has led to the identification of numerous gut bacteria from a range of different mosquito species (Gusmão *et al.* 2007, 2010; Lindh *et al.* 2005; Pidiyar *et al.* 2004; Rani *et al.* 2009), but it is a slow and expensive process so the number of sequences are usually relatively small. The advent of new sequencing technologies has both removed the need for cloning the PCR product and cut the cost of sequencing. This has led to 454 pyrosequencing being increasingly used to investigate microbial communities in other fields (Sogin *et al.* 2006; Roesch *et al.* 2007; Chandler *et al.* 2011; Huse *et al.* 2008; Bishop-Lilly *et al.* 2010; Wang *et al.* 2011).

In this study, we use 454 pyrosequencing to investigate the bacterial diversity in the guts of eight species of mosquitoes collected from the coastal region of Kenya. This allowed us to comprehensively catalogue the bacterial taxa present, and examine how the bacterial community varies in different mosquitoes. We found that mosquito gut typically has a very simple gut microbiota that is dominated by a single bacterial taxon. Although different mosquito species share remarkably similar gut bacteria, individuals in a population are extremely variable.

Materials and methods

Mosquito collection and identification

All mosquito samples were collected from towns and villages near Kilifi and Malindi on the Kenyan coast (Fig. S1, Supporting information). Collections were made in different localities in each area; Mbogolo in the Malindi district, and KEMRI, Mkwanjuni, Mnarani, Matsangoni and Jaribuni in the Kilifi district. BG-Sentinel traps (Biogents AG, Germany) or CDC light traps (Center for Disease Control, USA) were set to collect the mosquitoes. In general, *Anopheles* and *Mansonia* were collected from Mbogolo and Jaribuni, while *Aedes* and *Culex* were captured from the remaining sites. The mosquitoes were morphologically identified with the aid of taxonomic keys (Edwards 1941; Gillies & de Meillon 1968) and images from Walter Reed Biosystematics Unit (available from http://wrbu.si.edu/genera_mq.html). We later verified the mosquito identifications by amplifying the ribosomal internal transcribed spacer region-1 (ITS1) as described in von der Schulenburg *et al.* (2001), as different mosquito species produce different length of PCR products. PCR products were cleaned and sequenced with the BigDye Terminator kit (Perkin-Elmer Corporation, USA). Sequencing was performed at the Source BioScience Center, Cambridge, UK. Sequences were aligned and visually inspected in Sequencher v4.5 (Gene Codes Corporation). Resulting consensus sequences were searched against existing sequences in NCBI BLAST to confirm mosquito identification.

In total, 86 female mosquitoes distributed across four genera were analysed for their gut microflora. The number of amplified individuals for each species and location is shown on Table 1. We classified the sample sites into peri-urban and rural based on the type of buildings (concrete or mud) and infrastructure such as state of roads (tarred or untarred). Most of the Culicines were collected from Kilifi town and its environs, while all the Anophelines were collected from rural Mbogolo.

Dissection of mosquitoes

Females without blood-engorged abdomens were selected for dissection. Mosquitoes were surface sterilized prior to dissection; 10 min in dilute sodium hypochlorite, 1 min in 1× sterile phosphate buffered saline (PBS), 1 min in 70% ethanol and then a final wash in sterile 1× PBS. Dissections were performed under a stereomicroscope in a contained environment that was sterilized with 70% ethanol frequently to eliminate as much contamination as possible. Each mosquito gut was pulled out into a drop of sterile 1× PBS on a sterilized microscopic slide. Extracted guts were returned to a

Table 1 Mosquito samples collected in towns along the coast of Kenya

Mosquito species	District	Location	Habitat type	Number
Aedes aegypti	Kilifi	KEMRI	Peri-urban	3
Culex quinquefasciatus	Kilifi	KEMRI	Peri-urban	4
Aedes aegypti	Kilifi	Mnarani	Peri-urban	3
Aedes bromeliae	Kilifi	Mnarani	Peri-urban	10
Aedes aegypti	Kilifi	Mkwanjuni	Peri-urban	5
Aedes bromeliae	Kilifi	Mkwanjuni	Peri-urban	4
Aedes aegypti	Kilifi	Matsangoni	Rural	2
Anopheles gambiae	Kilifi	Jaribuni	Rural	1
Culex quinquefasciatus	Kilifi	Jaribuni	Rural	1
Anopheles gambiae	Malindi	Mbogolo	Rural	11
Anopheles funestus	Malindi	Mbogolo	Rural	11
Anopheles coustani	Malindi	Mbogolo	Rural	1
Mansonia africana	Malindi	Mbogolo	Rural	10
Mansonia uniformis	Malindi	Mbogolo	Rural	13
Culex quinquefasciatus	Malindi	Mbogolo	Rural	7

sterile 1.5 mL microcentrifuge tube containing about 500 μL absolute ethanol and stored till extraction.

DNA extraction

DNA from guts was extracted with QiAamp DNA Micro kit (Qiagen) according to the manufacturer's manual. All extractions were carried out under aseptic conditions; surfaces were cleaned with dilute sodium hypochlorite solution and 70% ethanol. To reduce contamination from bacteria in the surrounding air, all extractions were performed under a localized aseptic microenvironment provided by flame from a Bunsen burner. Microcentrifuge tubes for final DNA elution were irradiated with 200 mJ of ultraviolet light for 1 min in a UV Stratalinker 2400 (Stratagene Ltd., La Jolla, CA., USA) prior to use. A negative control, in which the extraction procedure was performed without adding any tissue, was included to check for contamination.

Primer design

We chose primers that amplified the V3 variable region of the 16S rRNA in Eubacteria, as this region is known to be informative in distinguishing bacterial species (Huse *et al.* 2008). The basic 16S primers, 338–358 F (5′ ACT CCT ACG GGA GGC AGC AGT 3′) and 683–700 R (5′ CGM ATT TCA CCK CTA CAC 3′) are highly conserved across the Eubacteria and amplify a region from position 359–682 (excluding primers) in the 16S rRNA of *Escherichia coli* (Wang & Qian 2009). To obtain a set of Fusion Primers (Roche), we added additional sequences required for Roche 454 Titanium Amplicon sequencing to the 5′ end of the primers. This also allowed multiplexing of the samples. Each complete Fusion Primer consisted of a 21-mer Primer A (5′ CGTATCGCCTCCCTCGCGCCA 3′) or Primer B (5′ CTATGCGCCTTGCCAGCCCGC 3′) followed by a 4-mer Key sequence (5′ TCAG 3′), a 10-mer Multiplex IDentifier (MID), and finally the 16S primer. The primers were HPLC purified. In total, we used 12 different MIDs for the forward primer and 12 different MIDs for the reverse primer (Table S1, Supporting information), which allowed us to multiplex up to 144 different samples in a single sequencing lane.

PCR and 454 parallel sequencing of gut bacteria

PCR amplification was performed with Phusion High-Fidelity DNA Polymerase (NEB, UK), following the manufacturer's recommendation for the reaction mix and cycle. Briefly, each 20 μL PCR reaction contained 4 μL of 5× buffer HF, 0.4 μL of 10 mM dNTP mix (Fermentas, UK), 0.4 μL of 20 μM forward and reverse Fusion Primer mix, 0.2 μL of 2 U/μL Phusion HF Polymerase, 1 μL of sample DNA and 14 μL of sterile water. All reactions were prepared under sterile conditions as described for DNA extraction earlier. Roughly equimolar concentrations of all positive samples were pooled into a single tube. The pooled sample was run on a 2% agarose gel and the resulting band excised and extracted from the gel using QIAquick gel extraction kit (Qiagen). The sample was then sequenced in both directions on an eighth of a Roche 454 FLX Genome Sequencer plate using Titanium Series reagents at the Department of Biochemistry Sequencing Facility, University of Cambridge.

Pre-processing of sequences

Sequencing on the GS FLX Automated Sequencer produced 49 576 sequences that had passed the machine's filter criteria. The average and median length of these sequences was 365.18 and 378.0 respectively, with a standard deviation of 48.79. The mean base quality score was 35.94. GS FLX reports base quality in Phred equivalent where a maximum score of 40 indicates a base calling accuracy of 99.99%, that is, a probability of one in 1000 that a base is incorrect (Margulies *et al.* 2005; Roche Applied Science 2009). As a quality control, we included in further analyses only sequences that were 340–400 bp long, had <20% ambiguous bases and an average quality score >25. The resulting sequences were grouped by their barcodes using the Geneious software (Drummond *et al.* 2001). During the barcode assignment, 4594 sequences did not have an exact match to our barcodes and so were not included. To remove chimeric sequences that arise during PCR, the

remaining 42 951 sequences were run through the chimera-slayer (Haas *et al.* 2011) program on Mothur (Schloss *et al.* 2009). To ensure we were only analysing bacterial sequences, a further 362 sequences with <75% similarity to any sequence in the SILVA–bacteria data set (Pruesse *et al.* 2007) were removed, and 14 sequences classified as chloroplast rDNA by Mothur (Pruesse *et al.* 2007) were also removed.

We proceeded to use the QIIME pipeline (Caporaso *et al.* 2010a) to organize the libraries by barcodes and align them. We aligned our sequences with the Python Nearest Alignment Space Termination Tool (PyNAST) (Caporaso *et al.* 2010b) using the Greengenes Core Set alignment as a template (DeSantis *et al.* 2006). We removed alignment columns where 95% of the positions were gaps. The aligned sequences were then assigned to Operational Taxonomic Units (OTUs), each with sequences sharing at least 97% similarity using the furthest-neighbour algorithm implemented in Mothur (Schloss *et al.* 2009). To reduce the number of sequences and enable faster analyses to be performed, representative sequences from each OTU were then selected and used in most analyses. Beta diversity estimates were made using UniFrac distances (Lozupone & Knight 2005) and are based on weighted Unifrac distances that have not been normalized unless otherwise stated. Other analyses such a rarefaction curves, heatmaps and statistical tests were carried out with custom scripts in R (R Development Core Team 2008).

Results

The mosquito gut has a low bacterial diversity

As bacterial species cannot be directly identified from our data, we classified the 33 757 sequences that passed our quality criteria into Operational Taxonomic Units (OTUs). Each OTU was defined by sequences with at least 97% nucleotide identity between them. In total, there were 789 OTUs, but only 53 of these ever exceeded a frequency of 1% in any of the guts we sampled. To estimate the species richness of the mosquito gut microbiota — the total number of OTUs present in a single gut — we used the Chao1 method (Chao & Lee 1992) to correct for our finite sample sizes. To assess the performance of this approach, we recalculated this statistic from different sized subsamples of the sequences from each gut and used these estimates to plot a rarefaction curve. As shown by the asymptotic curves in Fig. 1A, this analysis suggests that our sequencing depth was sufficient to obtain good estimates of the total species richness. Apart from two individuals with exceptionally high numbers of OTUs, we estimate that a mosquito gut contains ≈5–71 OTUs (Fig. 1A). The median number of OTUs in a mosquito gut is ≈42. The different species of mosquitoes had similar numbers of OTUs in their guts (Fig. 1B; $F_{7,69} = 1.50$, $P = 0.18$).

Despite a typical gut containing roughly 42 different OTUs (Fig. 1A), most of these are rare and the bacterial community is nearly always dominated by a small number of taxa. On average, the commonest OTU within a gut constituted 67% of all the bacteria sequenced from each sample, and the four most abundant OTUs together represented 90% of bacteria. This pattern of a few dominant OTUs within each gut can be clearly seen in the heatmap shown in Fig. 2A (see Table S2, Supporting information for OTU information). Therefore, the bacterial diversity – which reflects both the number and abundance of OTUs – is very low.

High variation within host species

There was extensive variation between individuals of the same host species in the composition of their gut microbiota. This is clear in the heatmaps shown in Fig. 2, which illustrate that it is common to find that an OTU or bacterial genus that makes up over 90% of the microbiota in one individual may be absent from the gut of another individual of the same species (Fig. 2). Certain OTUs were also exclusively found in a single individual within a host species (Fig. 2). Examples were observed in *Culex quinquefasciatus* (OTU 160), *Aedes aegypti* (OTU 500), *Anopheles gambiae* (OTU 215) and *Mansonia uniformis* (OTU 426).

To summarize the similarity of the gut microbiota in different individuals of the same species (β diversity), we calculated the weighted Unifrac distance between every pair of guts (Lozupone & Knight 2005). This statistic measures the distance between two communities by calculating the fraction of the branch length in a phylogenetic tree that leads to descendants in either, but not both, of the two communities. The weighted Unifrac distance that we used also accounts for the abundance of each bacterial taxa and is closely analogous to the fixation index F_{ST}. In this analysis, the distances were normalized, so a value of zero indicates that two guts have identical communities and a value of one that they have non-overlapping communities (i.e. when all the taxa are plotted on a phylogeny, no branches on the tree are shared). The average normalized Unifrac distances within species was 0.64, indicating that there is usually very little overlap in the composition of the gut microbiota of two individuals of the same species (Table 2).

Different host species have similar bacteria

Different species of mosquitoes have widely varying abilities to vector human parasites, so we were inter-

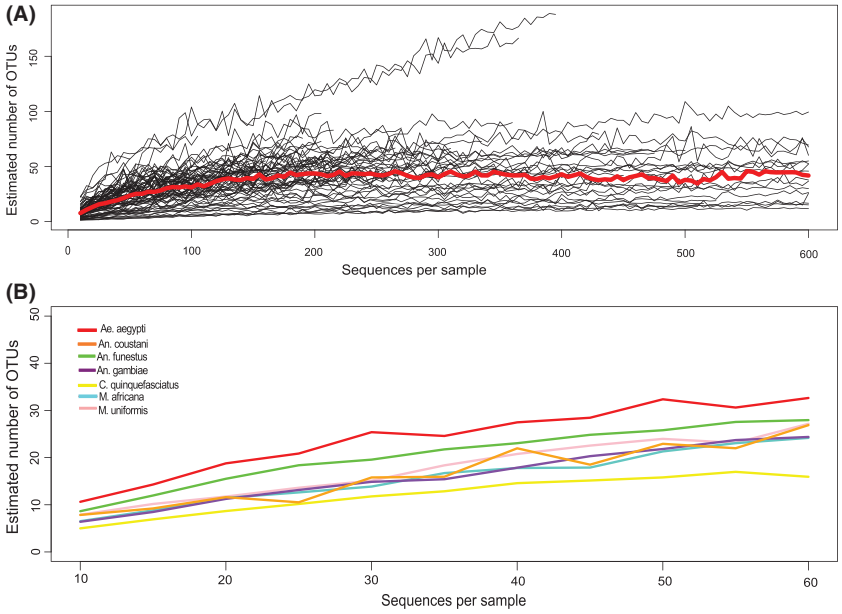

Fig. 1 The estimated number of bacterial Operational Taxonomic Units (OTUs) in mosquito guts. The number of OTUs in each gut was estimated using the Chao1 method (Chao & Lee 1992). The rarefaction curve was produced by randomly re-sampling different numbers of sequences from each individual (20 replicate samples/individual gut/sample size) and then calculating the mean of the 20 replicates. Panel A shows the estimated number of OTUs of the 86 mosquito guts re-sampled to a maximum depth of 600 sequences. The red line is the median number of OTUs at each level of sub-sampling. Panel B shows the mean of OTU estimates for the guts from each mosquito species. Note that there was only a single individual of *Anopheles coustani*.

ested in whether each species had a unique gut microbiota that could be influencing their vector competence. To do this, we examined how the total bacterial diversity was partitioned among individuals within each species and between species using the weighted Unifrac statistic described earlier. We found that while different individuals did have significantly different gut microbiota (Mantel test on matrix of weighted Unifrac distances: $r = -0.07$, $P < 0.001$), only 7% of the variation was explained by between-species differences. Therefore, two individuals from the same mosquito species will typically have almost as great a difference in their gut microbiota as two individuals from different species. After taking mosquito species into account, sampling location had no significant correlation with the species composition of bacteria found in guts (Partial Mantel test on matrix of weighted Unifrac distances: $r = -0.01$, $P = 0.33$).

To visualize these differences between species, we used the matrix of weighted UniFrac distances to construct a UPGMA tree (Lozupone *et al.* 2007; Lozupone & Knight 2005). From this analysis, it is clear that it is normal for individual mosquitoes to have gut microbiota that are more similar to individuals in other species than individuals from the same species (Fig. 3). The same pattern is evident if the same data is used to create principal components plots (Fig. 4). Despite this, there is clearly a tendency for certain species to cluster together. For example, *Culex* and *Mansonia* species tend to have similar gut bacteria, as does *Ae. aegypti* and the *Anopheles* species (Figs 3 and 4).

Classification of bacterial OTUs

Different bacterial taxa can have very different effects on the vectorial capacity of mosquitoes, so we classified our OTUs to the level of genus (Figs 2B and 5). To do this, we compared our filtered sequence reads to 16S rRNA sequences in the Ribosomal Database Project (RDP II) Library using the Bayesian approach that is implemented by the RDP Classifier (Table S3, Supporting information) (Wang *et al.* 2007). This classification resulted in 144 unique bacterial genera that were mainly composed of four abundant classes of bacteria. The Gram-negative Gammaproteobacteria, Alphaproteobacteria and Flavobacteria represented 62.3%, 18.3% and 11.6% of these classified bacteria, respectively. The Gram-positive Bacilli constituted 3.8% of the bacteria. 17.5% of the bacteria could not be classified below the level of class (Fig. 5). Nearly half of all the classified bacteria belonged to two genera (Figs 2B and 5) – *Aeromonas* (38.7%) and *Asaia* (13.2%). The next most abun-

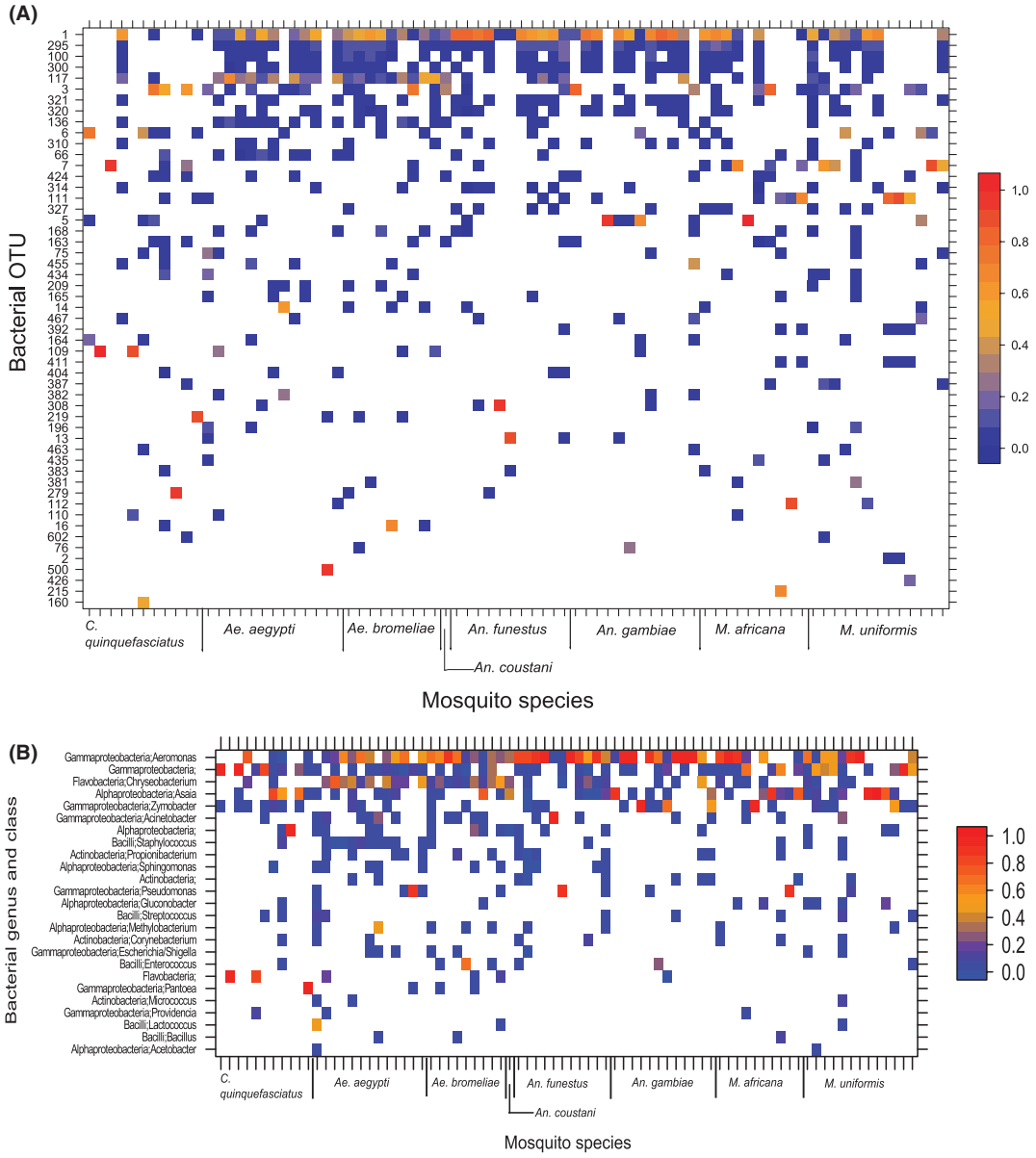

Fig. 2 The frequency of different bacterial Operational Taxonomic Units (OTUs) (panel A) and genera (panel B) in the guts of individual mosquitoes. Each column is a different individual, and each row is a different OTU or genus. The colour represents the proportion of sequence reads from a given OTU or genus in that mosquito. White spaces are OTUs found at a frequency of <1%. The OTUs are arranged so the most frequently occurring are at the top of the figure, and only individuals with at least 20 sequence reads are included. In panel B, the absence of a genus name indicates that the sequences could only be classified to the level of Class.

dant genus was *Chyseobacterium* (9.1%) followed by *Zymobacter* (6.0%).

Other bacterial taxa tended to have a much more patchy distribution (Fig. 2B). For example, unclassified Flavobacteria dominate in two individuals of *C. quinquefasciatus* (82% and 99% of sequences), but are rare or absent in the rest of our samples. *Pantoea* (Enterobacteriaceae) was dominant in a single individual of *C. quinquefasciatus* (94%), while *Pseudomonas* was at frequencies above 80% in single individuals of *Ae. aegypti*, *An. funestus* and *M. africana*. The distribution also suggests there are no host-specific bacterial genera. As was the case for the analysis of OTUs, most of the variation in the bacterial taxa was between individuals within a species rather than between species (Fig. 2B). Only two of the bacterial genera showed significant variation in abundance in the different mosquito species – *Aeromonas*, which varied from 49.8% in *An. funestus* to 14.9% in *M. uniformis* (Kruskal–Wallis test: $\chi^2 = 22.53$, d.f. = 7, $P = 0.002$),

Table 2 Variation in the gut microbiota of individuals of the same species. The distances between individuals of the same species are normalized, weighted Unifrac distances. A value of 0 indicates identical bacterial communities, and a value of 1 indicates no phylogenetic overlap between the communities

Host species	Mean distance
Culex quinquefasciatus	0.63
Mansonia uniformis	0.64
Mansonia africana	0.68
Anopheles gambiae	0.74
Anopheles funestus	0.72
Aedes aegypti	0.43
Aedes bromeliae	0.66

and *Chryseobacterium* (Kruskal–Wallis test: $\chi^2 = 33.6$, d.f. $= 7$, $P = 2 \times 10^{-5}$).

Discussion

We have provided a comprehensive, unbiased, culture-independent study of the bacterial community in the guts of eight mosquito species sampled from natural populations. We found that there is generally a very low bacterial diversity, with a single OTU typically making up two-thirds of all the bacteria (the divergence between our OTUs is similar to that between many bacterial species, although some OTUs may be strains of the same species (Schleifer 2008)). However, there are also many other rarer bacteria, with a typical gut containing 42 bacterial OTUs. Between individual mosquitoes of the same species, there is enormous variation in the bacterial taxa present, but there were few consistent differences between the different mosquito species in the composition of their gut microbiota.

The variation between individuals in the composition of their gut microbiota may affect the vector competence of mosquitoes. Several studies have found that a range of Gram-negative gut bacteria inhibit the development of *Plasmodium*, while Gram-positive bacteria do not (Cirimotich *et al.* 2011; Gonzalez-Ceron *et al.* 2003; Pumpuni *et al.* 1993). Furthermore, different Gram-negative bacteria have varying effects on *Plasmodium* (Cirimotich *et al.* 2011; Gonzalez-Ceron *et al.* 2003; Jadin 1967). Some of the variation may be explained by the differences in the production of certain metabolites. For example, the red pigment prodigiosin, which is produced by some Gram-negative bacteria, has been shown to be effective against *Plasmodium* (Kim *et al.* 1999; Isaka *et al.* 2002; Lazaro *et al.* 2002). In *Anopheles stephensi*, *Klebsiella* blocked the development of *Plasmodium berghei*, whereas *Pseudomonas* did not (Jadin 1967). In the same mosquito species, the two bacteria genera had opposite effects on *P. falciparum* (Jadin 1967). Therefore, the differences in the gut microbiota we have observed between individuals within a host species could be causing variation in vector competence. They may also be affecting other aspects of host biology, such as is observed in *Drosophila melanogaster* where flies prefer to mate with individuals that have a similar bacteria community in their guts (Sharon *et al.* 2010).

The variation that we have observed may reflect differences in the bacteria that mosquitoes have acquired from the environment. Mosquitoes such as *Anopheles*, *Aedes* and *Culex* prefer laying their eggs in water that contains bacteria (Lindh *et al.* 2008; Pavlovich & Rockett 2000; Rockett 1987), and midgut bacteria acquired from the larval environment can then be transmitted transstadially to the adult gut (Jadin *et al.* 1966; Pumpuni *et al.* 1996; Briones *et al.* 2008). It is also possible for adult mosquitoes to acquire bacteria from their breeding water while they emerge from their pupal cases (Lindh *et al.* 2008). Bacteria acquired this way can then be horizontally transferred between individuals through deposition of bacteria back into laying water (Lindh *et al.* 2008) or via common feeding sites.

Host diet also shapes the gut microbiome. Differences in the gut microbiota of several species of *Drosophila* is strongly influenced by diet (Chandler *et al.* 2011), and it is possible that the lack of between-species variation we observed in the gut microbiota is because these mosquitoes tend to have rather similar diets, feeding on microbes as larvae, and blood and nectar as adults. Our samples may have included individuals that had blood fed and individuals that had not. This could cause between-individual variation, as both sugar- and blood feeding changes bacterial abundance in mosquito midguts (Demaio *et al.* 1996; Gusmão *et al.* 2010; Wang *et al.* 2011). Acetic acid bacteria, for example, are associated with many insects that have a sugar-based diet (Ashbolt & Inkerman 1990; Mohr & Tebbe 2006; Corby-Harris *et al.* 2007; Crotti *et al.* 2009). Blood feeding, in particular, triggers the proliferation of bacteria (Gusmão *et al.* 2010) and, certain bacterial taxa show more increase than others. In *Anopheles gambiae*, older adults that have had blood meals increase Proteobacteria and decrease the bacterial species richness in the gut (Wang *et al.* 2011). Newly emerged adults that have not fed generally have higher species richness than we observed (Wang *et al.* 2011), which suggests that we may have sampled older mosquitoes that have already fed. Although we selected individuals that were not engorged with blood, we have no knowledge of prior feeding patterns of our samples and how old they could be since emergence.

A less well understood influence on the bacterial community is the genetic background of the insect. Evidence for its potential role comes from *Drosophila*,

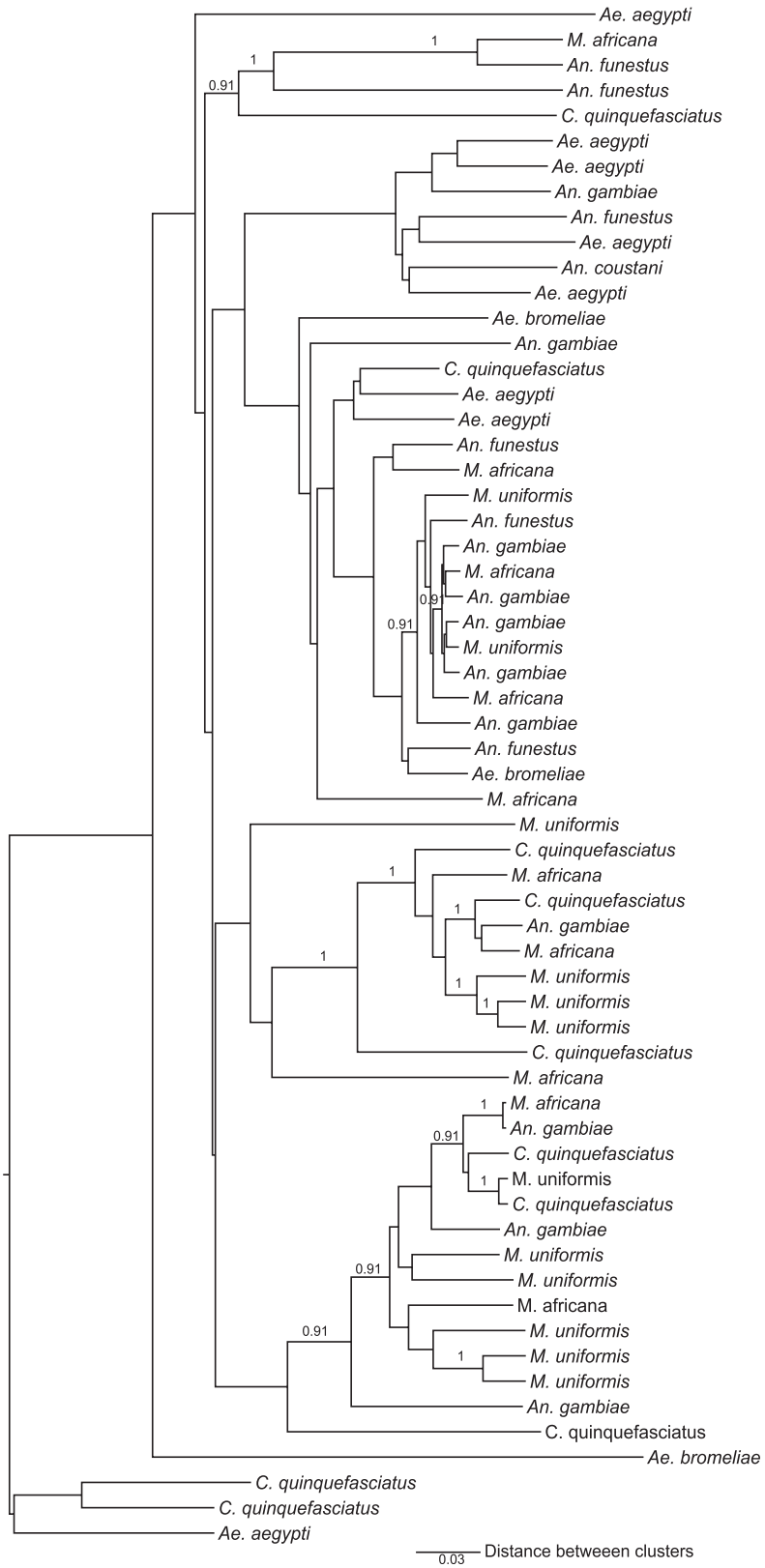

Fig. 3 UPGMA tree showing clustering of the bacterial communities in the mosquito guts. Weighted (quantitative) classification was used. Jackknifed support values above 0.90 are shown.

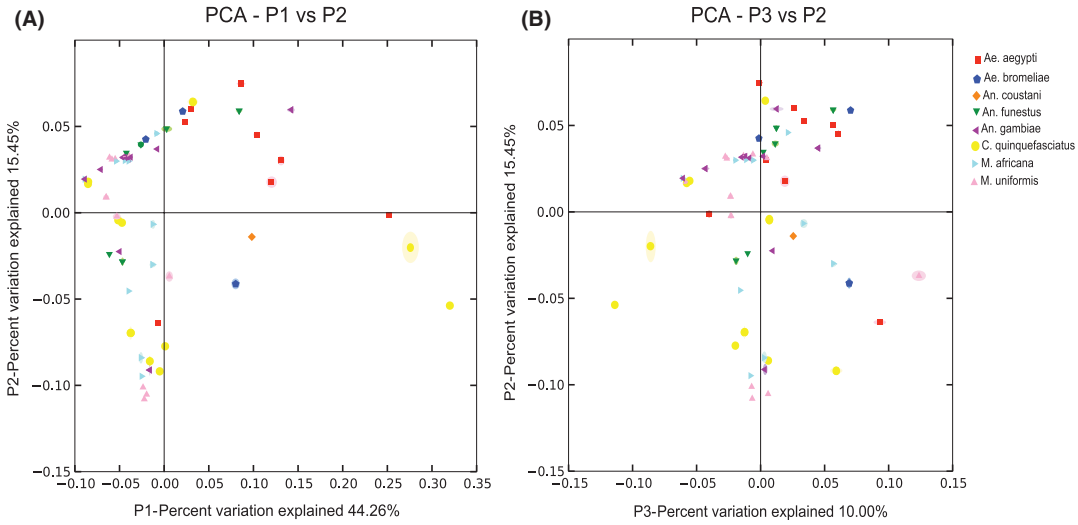

Fig. 4 Principal component (PC) plots of gut bacteria diversity. Coloured dots represent mosquito samples. Lack of ellipsoids around dots indicates strong Jackknife support values.

where the gene *Caudal* maintains immune system homoeostasis, and knock-down of the gene alters the composition of the gut microbial community, resulting in high mortality (Ryu *et al.* 2008). However, as the differences between individuals of the same species are far greater than between different species, this is likely to be a relatively unimportant factor.

Our deep sequencing approach has allowed us to catalogue the taxonomic diversity of the mosquito gut microbiota in great detail. In total, we found 22 genera that occur at a frequency of more than 1% in at least one of the individual guts. Consistent with previous studies, we found that the mosquito gut is dominated by Gram-negative bacteria (Cirimotich *et al.* 2011; Demaio *et al.* 1996; Dong *et al.* 2009; Lindh *et al.* 2005; Straif *et al.* 1998), with Proteobacteria and Bacteroidetes constituting more than 90% of the community. As Gram-negative bacteria tend to offer greater protection against *Plasmodium* than Gram-positive bacteria (Cirimotich *et al.* 2011; Gonzalez-Ceron *et al.* 2003; Pumpuni *et al.* 1993, 1996), this suggests that the gut microbiota may be an important factor reducing rates of disease transmission.

Aeromonas spp. was the commonest bacteria, representing on average over a third of the gut microbiota. *Aeromonas spp.* are common in insects, having been previously reported in house flies (Nayduch *et al.* 2005), tsetse flies (Geiger *et al.* 2011) and mosquitoes (Djadid *et al.* 2011; Pidiyar *et al.* 2004), with *Aeromonas culiciola* being the most abundant gut bacterium in *Culex quinquefasciatus* (Pidiyar *et al.* 2002). *Aeromonas* is also commonly isolated from breeding water of mosquitoes (Smith *et al.* 1998), suggesting that mosquitoes are ingesting these bacteria as larvae. Transstadial transfer from larvae to the adult gut is possible. Despite a reduction in bacterial numbers in the adult gut (Chavshin *et al.* 2012), *Aeromonas* rapidly proliferates following a blood meal (Pidiyar *et al.* 2002).

The second most abundant genus was *Asaia*, which is found in all the mosquito species we sampled at an average frequency of 13%. *Asaia* is an acetic acid bacterium that has been found in the midgut, salivary glands and reproductive organs of *An. stephensi* and *An. gambiae*, two species of mosquitoes that transmit malaria (Favia *et al.* 2007; Damiani *et al.* 2010). The localization of *Asaia* in these tissues means that they may play important roles in interacting with parasites. Furthermore, the bacterium is not only transmitted horizontally when mosquitoes feed together, but it is also transmitted sexually, maternally and paternally, so it can form stable associations across multiple generations (Damiani *et al.* 2008). Unusual for a vertically transmitted symbiont of insects, it can also be cultured, transformed and easily moved between host species (Crotti *et al.* 2009), making it an excellent candidate for expressing anti-parasite proteins in natural populations, an approach called paratransgenesis (Favia *et al.* 2007; Damiani *et al.* 2010). Our study is the first report of natural association of *Asaia* with *Mansonia uniformis*, *M. africana*, *Aedes bromeliae* and *An. coustani*. *Mansonia uniformis* is a competent vector of filariasis (Nelson 1959; Ramalingam 1968; Wharton 1962), and *Ae. bromeliae* transmits yellow fever (Huang 1986). Although *An. coustani* is previously recognized to have a zoophilic behaviour, it has recently been reported to also possess anthrophilic tendencies implicating the species in the potential transmission of human malaria (Fornadel *et al.* 2011). Our results suggest that *Asaia*

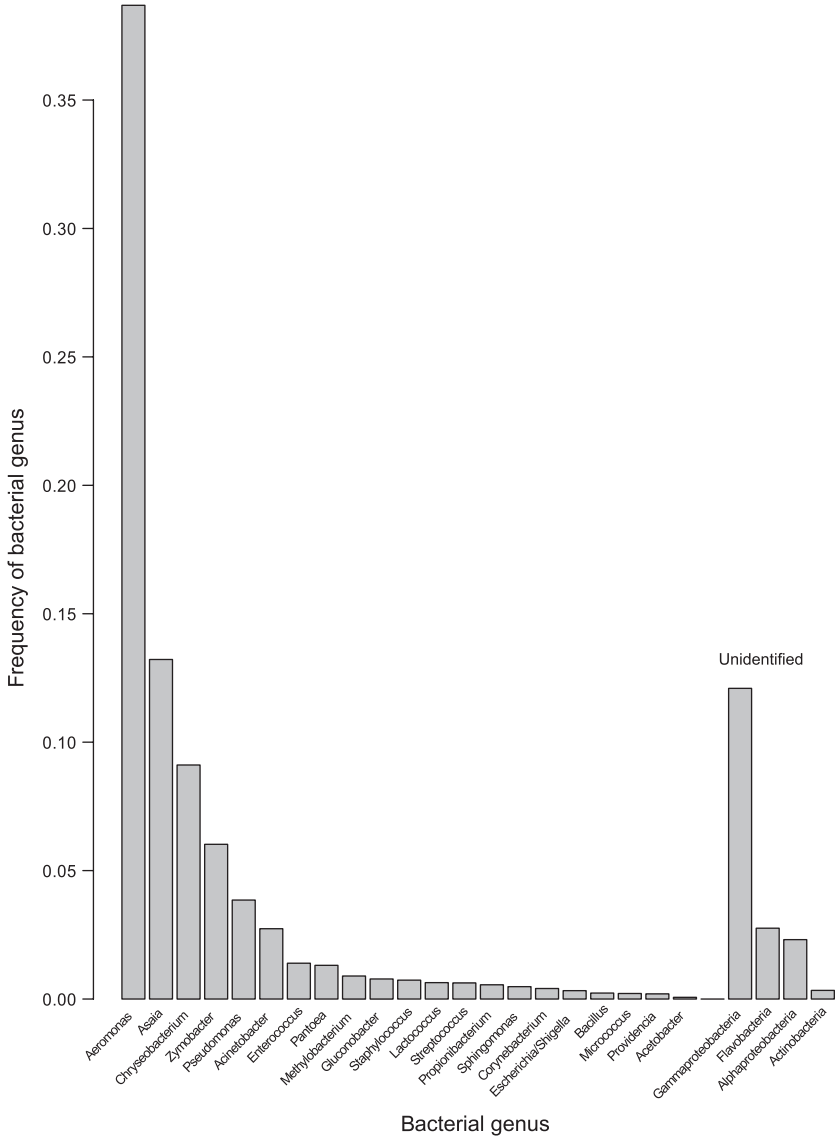

Fig. 5 The mean frequency of bacterial genera in mosquito guts. Only bacteria that exceed 1% frequency in at least one individual are included. The mean frequencies are calculated with each individual gut being weighted equally.

can infect field populations of most mosquito disease vectors.

Chryseobacterium was the third most abundant genus and was particularly frequent in *Aedes* mosquitoes. Although *Chryseobacterium meningosepticum* was found in all individuals of *An. gambiae* tested by Dong *et al.* (2009), little is known about its effects on insects or disease transmission.

Pseudomonas has been reported to be common in mosquito guts (Jadin *et al.* 1966; Rani *et al.* 2009), but we found it had a very patchy distribution with only a few individuals infected at frequencies above 1%. Both positive and negative effects of *Pseudomonas* on *Plasmodium* have been reported (Jadin *et al.* 1966; Straif *et al.* 1998), so the heterogeneity in infection rates may contribute to variation in disease transmission. *Pseudomonas* proliferates after a blood meal, and it has been suggested that it may be important in coping with oxidative stress after blood feeding (Wang *et al.* 2011).

In *Anopheles* mosquitoes, bacteria in the genus *Enterobacter* can dramatically reduce the intensity of *Plasmodium* infection (Cirimotich *et al.* 2011; Straif *et al.* 1998) because of the production of reactive oxygen species that affect the development of oocysts from ookinetes (Cirimotich *et al.* 2011). They have also been reported to be common in blood-fed *Ae. aegypti* (Gaio *et al.* 2011) where they have haemolytic activity that is important in digestion. Despite these bacteria having been isolated from African mosquitoes in the past, we found *Enterob-*

acter were rare in our sample, never occurring at a frequency over 1% in any of the guts.

The bacteria we have identified might not only be important in affecting natural rates of disease transmission, but could also be exploited to manipulate disease transmission in the wild. It may be possible to infect wild mosquito populations with bacteria that either naturally confer resistance to human parasites, or to exploit them for paratransgenesis – the concept of using insect symbionts to drive anti-parasitic factors through populations (Beard *et al.* 2002). Before this can be attempted, we need a greater understanding of the transmission of these bacteria in nature, and how bacteria released into the environment can compete with the natural symbionts of mosquitoes.

In summary, deep sequencing has allowed us to provide a comprehensive catalogue of the bacteria that naturally inhabit mosquito guts in this population, expanding the range of both hosts and bacteria that have been studied. We found that the mosquito microbiota has a very low bacterial diversity within an individual, but much greater variation across different individuals. Understanding the implications of this variation for disease transmission promises to be a fertile field for future research.

Acknowledgements

This work was funded by a Royal Society Research Fellowship to FMJ, a Darwin Trust of Edinburgh studentship to JOP and an MRC studentship to WJP. The field work was supported by the following grants to JOP: Genetics Society Heredity Fieldwork Grant, Royal Entomological Outreach Grant and Chibnall Travel Funds. We want to thank staff of the Vector Biology Group, KEMRI-Wellcome Trust Research Programme, Kilifi, for their technical assistance and advice during the field collection and laboratory work in Kenya, especially, to Mr. Festus Nah for his immense help in mosquito sampling.

References

Al-Olayan EM, Williams GT, Hurd H (2002) Apoptosis in the malaria protozoan, *Plasmodium berghei*: a possible mechanism for limiting intensity of infection in the mosquito. *International Journal for Parasitology*, **32**, 1133–1143.

Ashbolt NJ, Inkerman PA (1990) Acetic acid bacterial biota of the pink sugar cane mealybug, *Saccharococcus sacchari*, and its environs. *Applied and Environmental Microbiology*, **56**, 707–712.

Azambuja P, Garcia ES, Ratcliffe NA (2005) Gut microbiota and parasite transmission by insect vectors. *Trends in Parasitology*, **21**, 568–572.

Beard CB, Cordon-Rosales C, Durvasula RV (2002) Bacterial symbionts of the Triatominae and their potential use in control of Chagas disease transmission. *Annual Review of Entomology.*, **47**, 123–141.

Bishop-Lilly KA, Turell MJ, Willner KM, *et al.* (2010) Arbovirus detection in insect vectors by rapid, high-throughput pyrosequencing. *PLoS Neglected Tropical Diseases*, **4**, e878.

Blandin S, Shiao S-H, Moita LF, *et al.* (2004) Complement-like protein TEP1 is a determinant of vectorial capacity in the malaria vector *Anopheles gambiae*. *Cell*, **116**, 661–670.

Briones AM, Shililu J, Githure J, *et al.* (2008) *Thorsellia anophelis* is the dominant bacterium in a Kenyan population of adult *Anopheles gambiae* mosquitoes. *ISME Journal*, **2**, 74–82.

Caporaso JG, Kuczynski J, Stombaugh J, *et al.* (2010a) QIIME allows analysis of high-throughput community sequencing data. *Nature Methods*, **7**, 335–336.

Caporaso JG, Kyle B, Bushman FD, *et al.* (2010b) PyNAST: a flexible tool for aligning sequences to a template alignment. *Bioinformatics*, **26**, 266–267.

Chandler JA, Lang JM, Bhatnagar S, *et al.* (2011) Bacterial communities of diverse *Drosophila* species: ecological context of a host–microbe model system. *PLoS Genetics*, **7**, e1002272.

Chao A, Lee SM (1992) Estimating the number of classes via sample coverage. *Journal of the American Statistical Association*, **87**, 210–217.

Chavshin AR, Oshaghi MA, Vatandoost H, *et al.* (2012) Identification of bacterial microflora in the midgut of the larvae and adult of wild caught *Anopheles stephensi*: a step toward finding suitable paratransgenesis candidates. *Acta Tropica*, **121**, 129–134.

Cirimotich CM, Dong Y, Clayton AM, *et al.* (2011) Natural microbe-mediated refractoriness to *Plasmodium* infection in *Anopheles gambiae*. *Science*, **332**, 855–858.

Corby-Harris V, Pontaroli AC, Shimkets LJ, *et al.* (2007) Geographical distribution and diversity of bacteria associated with natural populations of *Drosophila melanogaster*. *Applied and Environmental Microbiology*, **73**, 3470–3479.

Crotti E, Damiani C, Pajoro M, *et al.* (2009) *Asaia*, a versatile acetic acid bacterial symbiont, capable of cross-colonizing insects of phylogenetically distant genera and orders. *Environmental Microbiology*, **11**, 3252–3264.

Damiani C, Ricci I, Crotti E, *et al.* (2008) Paternal transmission of symbiotic bacteria in malaria vectors. *Current Biology*, **18**, 1087–1088.

Damiani C, Ricci I, Crotti E, *et al.* (2010) Mosquito-bacteria symbiosis: the case of *Anopheles gambiae* and *Asaia*. *Microbial Ecology*, **60**, 644–654.

Demaio J, Pumpuni CB, Kent M, Beier JC (1996) The midgut bacterial flora of wild *Aedes triseriatus*, *Culex pipiens*, and *Psorophora columbiae* mosquitoes. *The American Journal of Tropical Medicine and Hygiene*, **54**, 219–223.

DeSantis TZ, Hugenholtz N, Larsen M, *et al.* (2006) Greengenes, a chimer-checked 16S rRNA gene database and workbench compatible with ARB. *Applied and Environmental Microbiology*, **72**, 5069–5072.

Djadid DN, Jazayeri H, Raz A, *et al.* (2011) Identification of the midgut microbiota of *An. stephensi* and *An. maculipennis* for their application as a paratransgenic tool against malaria. *PLoS ONE*, **6**, e28484.

Dong Y, Manfredini F, Dimopoulos G (2009) Implication of the mosquito midgut microbiota in the defense against malaria parasites. *PLoS Pathogens*, **5**, e1000423.

Drummond AJ, Ashton B, Buxton S, *et al.* (2001) Geneious v5.4. Available from http://www.geneious.com/.

Edwards FW (1941) Mosquitoes of the Ethiopian Region III- Culicine adults and pupae. The British Museum of Natural History, London, pp 499.

Erickson SM, Xi Z, Mayhew GF, *et al.* (2009) Mosquito infection responses to developing filarial worms. *PLoS Neglected Tropical Diseases*, **3**, e529.

Favia G, Ricci I, Damiani C, *et al.* (2007) Bacteria of the genus *Asaia* stably associate with *Anopheles stephensi*, an Asian malarial mosquito vector. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 9047–9051.

Fornadel CM, Norris LC, Franco V, Norris DE (2011) Unexpected anthropophily in the potential secondary malaria vectors *Anopheles coustani s.l.* and *Anopheles squamosus* in Macha, Zambia. *Vector-Borne and Zoonotic Diseases*, **11**, 1173–1179.

Gaio ADO, Gusmão DS, Santos AV, *et al.* (2011) Contribution of midgut bacteria to blood digestion and egg production in *Aedes aegypti* (Diptera: Culicidae) (L.). *Parasites & Vectors*, **4**, 105.

Geiger A, Fardeau M-L, Njiokou F, *et al.* (2011) Bacterial diversity associated with populations of *Glossina spp.* from Cameroon and distribution within the Campo sleeping sickness focus. *Microbial Ecology*, **62**, 632–643.

Gene Codes Corporation. Sequencher® version 4.5. Sequence analysis software.

Gillies MT, de Meillon B (1968) The Anophelinae of Africa South of the Sahara (Ethipian Zoogeographical Region). Publications of the South African Institute for Medical Research.

Gonzalez-Ceron L, Santillan F, Rodriguez MH, *et al.* (2003) Bacteria in midguts of field-collected *Anopheles albimanus* block *Plasmodium* vivax sporogonic development. *Journal of Medical Entomology*, **40**, 371–374.

Gusmão DS, Santos AV, Marini DC, *et al.* (2007) First isolation of microorganisms from the gut diverticulum of *Aedes aegypti* (Diptera: Culicidae): new perspectives for an insect-bacteria association. *Memórias do Instituto Oswaldo Cruz*, **102**, 919–924.

Gusmão DS, Santos AV, Marini DC, *et al.* (2010) Culture-dependent and culture-independent characterization of microorganisms associated with *Aedes aegypti* (Diptera: Culicidae) (L.) and dynamics of bacterial colonization in the midgut. *Acta Tropica*, **115**, 275–281.

Haas BJ, Gevers D, Earl AM, *et al.* (2011) Chimeric 16S rRNA sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons. *Genome Research*, **21**, 494–504.

Huang Y-M (1986) *Aedes* (Stegomyia) *bromeliae*: the Yellow fever (Diptera: Culicidae) vector in East Africa. *Journal of Medical Entomology*, **23**, 196–200.

Huse SM, Dethlefsen L, Huber JA, *et al.* (2008) Exploring microbial diversity and taxonomy using SSU rRNA hypervariable tag sequencing. *PLoS Genetics*, **4**, e1000255.

Isaka M, Jaturapat A, Kramyu J, *et al.* (2002) Potent in vitro antimalarial activity of metacycloprodigiosin isolated from *Streptomyces spectabilis* BCC 4785. *Antimicrobial Agents and Chemotherapy*, **46**, 1112–1113.

Jadin J (1967) Du role des bacteries dans les tubes digestif des insects vecteurs des plasmodidae et des trypanosomidae. *Annales de la Société belge de médecine tropicale*, **47**, 331–342.

Jadin J, Vincke IH, Dunjic A, *et al.* (1966) Role of Pseudomonas in the sporogenesis of the hematozoon of malaria in the mosquito. *Bulletin de la Société de pathologie exotique et de ses filiales*, **59**, 514–525.

Kim HS, Hayashi M, Shibata Y, *et al.* (1999) Cycloprodigiosin hydrochloride obtained from *Pseudoalteromonas denitrificans* is a potent anti-malarial agent. *Biological & Pharmaceutical Bulletin*, **22**, 532–534.

Lazaro JEH, Nitcheu J, Predicala RZ, *et al.* (2002) Heptyl prodigiosin, a bacterial metabolite, is antimalarial in vivo and non-mutagenic in vitro. *Journal of Natural Toxins*, **11**, 367–377.

Lindh JM, Terenius O, Faye I (2005) 16S rRNA gene-based identification of midgut bacteria from field-caught *Anopheles gambiae sensu lato* and *A. funestus* mosquitoes reveals new species related to known insect symbionts. *Applied and Environmental Microbiology*, **71**, 7217–7223.

Lindh JM, Borg-Karlson A-K, Faye I (2008) Transstadial and horizontal transfer of bacteria within a colony of *Anopheles gambiae* (Diptera: Culicidae) and oviposition response to bacteria-containing water. *Acta Tropica*, **107**, 242–250.

Lozupone C, Knight R (2005) UniFrac: a New phylogenetic method for comparing microbial communities. *Applied and Environmental Microbiology*, **71**, 8228–8235.

Lozupone C, Hamady M, Kelley ST, Knight R (2007) Quantitative and qualitative β diversity measures lead to different insights into factors that structure microbial communities. *Applied and Environmental Microbiology*, **73**, 1576–1585.

Margulies M, Egholm M, Altman WE, *et al.* (2005) Genome sequencing in microfabricated high -density picolitre reactors. *Nature*, **437**, 376–380.

Michalski ML, Erickson SM, Bartholomay LC, Christensen BM (2010) Midgut barrier imparts selective resistance to filarial worm infection in *Culex pipiens pipiens*. *PLoS Neglected Tropical Diseases*, **4**, e875.

Michel K, Kafatos FC (2005) Mosquito immunity against *Plasmodium*. *Insect Biochemistry and Molecular Biology*, **35**, 677–689.

Mohr KI, Tebbe CC (2006) Diversity and phylotype consistency of bacteria in the guts of three bee species (Apoidea) at an oilseed rape field. *Environmental Microbiology*, **8**, 258–272.

Müller GC, Beier JC, Traore SF, *et al.* (2010) Successful field trial of attractive toxic sugar bait (ATSB) plant-spraying methods against malaria vectors in the *Anopheles gambiae* complex in Mali, West Africa. *Malaria Journal*, **9**, 210.

Nayduch D, Noblet GP, Stutzenberger F (2005) Fate of bacteria, *Aeromonas caviae*, in the midgut of the housefly, *Musca domestica*. *Invertebrate Biology*, **124**, 74–78.

Nelson GS (1959) The identification of infective filarial larvae in mosquitoes: with a note on the species found in "wild" mosquitoes on the Kenya coast. *Journal of Helminthology*, **33**, 233–256.

Osta MA, Christophides GK, Kafatos FC (2004) Effects of mosquito genes on *Plasmodium* development. *Science (New York, N.Y.)*, **303**, 2030–2032.

Pavlovich SG, Rockett CL (2000) Color, bacteria and mosquito eggs as ovipositional attractants for *Aedes aegypti* and *Aedes albopictus* (Diptera: Culicidae). *Great Lake Entomologist*, **33**, 141–153.

Pidiyar V, Kaznowski A, Narayan NB, *et al.* (2002) *Aeromonas culicicola sp. nov* ., from the midgut of *Culex quinquefasciatus*. *International Journal of Systematic and Evolutionary Microbiology*, **52**, 1723–1728.

Pidiyar VJ, Jangid K, Patole MS, Shouche YS (2004) Studies on cultured and uncultured microbiota of wild *Culex quinquefasciatus* mosquito midgut based on 16S ribosomal RNA gene analysis. *The American Journal of Tropical Medicine and Hygiene*, **70**, 597–603.

Pruesse E, Quast C, Knittel K, *et al.* (2007) SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Research*, **35**, 7188–7196.

Pumpuni CB, Beier MS, Nataro JP, *et al.* (1993) *Plasmodium falciparum*: inhibition of sprogonic development in *Anopheles stephensi* by Gram-negative bacteria. *Experimental Parasitology*, **77**, 195–199.

Pumpuni CB, Demaio J, Kent M, *et al.* (1996) Bacterial population dynamics in three Anopheline species: the impact on *Plasmodium* sporogonic development. *American Journal of Tropical Medicine and Hygiene*, **54**, 214–218.

R Development Core Team (2008) R: A Language and Environment for Statistical Computing.

Ramalingam S (1968) The epidemiology of filarial transmission in Samoa and Tonga. *Annals of Tropical Medicine and Parasitology*, **62**, 305–324.

Rani A, Sharma A, Rajagopal R, *et al.* (2009) Bacterial diversity analysis of larvae and adult midgut microflora using culture-dependent and culture-independent methods in lab-reared and field-collected *Anopheles stephensi*-an Asian malarial vector. *BMC Microbiology*, **9**, 96.

Richman AM, Dimopoulos G, Seeley D, Kafatos FC (1997) *Plasmodium* activates the innate immune response of *Anopheles gambiae* mosquitoes. *The EMBO Journal*, **16**, 6114–6119.

Roche Applied Science (2009) Genome Sequrncer FLX System Software Manual, version 2.3. 1–42.

Rockett CL (1987) Bacteria as ovipositional attractants for *Culex pipiens* (Diptera: Culicidae). *Great Lake Entomologist*, **20**, 151–155.

Roesch LFW, Fulthorpe RR, Riva A, *et al.* (2007) Pyrosequencing enumerates and contrasts soil microbial diversity. *ISME Journal*, **1**, 283–290.

Ryu JH, Kim SH, Lee HY, *et al.* (2008) Innate immune homeostasis by the homeobox gene Caudal and commensal-gut mutualism in *Drosophila*. *Science*, **319**, 777–782.

Schleifer KH (2008) Classification of bacteria and archaea: past, present and future. *Systematic and Applied Microbiology*, **32**, 533–542.

Schloss PD, Westcott SL, Ryabin T, *et al.* (2009) Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology*, **75**, 7537–7541.

von der Schulenburg JH, Habig M, Sloggett JJ, *et al.* (2001) Incidence of male-killing *Rickettsia spp.* (alpha-proteobacteria) in the ten-spot ladybird beetle *Adalia decempunctata* L. (Coleoptera: Coccinellidae). *Applied and Environmental Microbiology*, **67**, 270–277.

Sharon G, Segal D, Ringo JM, *et al.* (2010) Commensal bacteria play a role in mating preference of *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences of the United States of America*, **107**, 20051–20056.

Smith TW, Walker ED, Kaufman MG (1998) Bacterial density and survey of cultivable heterotrophs in the surface water of a freshwater marsh habitat of *Anopheles quadrimaculatus* larvae (Diptera: Culicidae). *Journal of the American Mosquito Control Association*, **14**, 72–77.

Sogin ML, Morrison HG, Huber JA, *et al.* (2006) Microbial diversity in the deep sea and the underexplored "rare biosphere". *Proceedings of the National Academy of Sciences*, **103**, 12115–12120.

Straif SC, Mbogo CNM, Toure AM, *et al.* (1998) Midgut bacteria in *Anopheles gambiae* and *An-funestus* (Diptera: Culicidae) from Kenya and Mali. *Journal of Medical Entomology*, **35**, 222–226.

Vlachou D, Schlegelmilch T, Christophides GK, Kafatos FC (2005) Functional genomic analysis of midgut epithelial responses in *Anopheles* during *Plasmodium* invasion. *Current Biology: CB*, **15**, 1185–1195.

Wang Y, Qian P-Y (2009) Conservative fragments in bacterial 16S rRNA genes and primer design for 16S ribosomal DNA amplicons in metagenomic studies. *PLoS ONE*, **4**, e7401.

Wang Q, Garrity GM, Tiedje JM, Cole JR (2007) Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and Environmental Microbiology*, **73**, 5261–5267.

Wang Y, Gilbreath TM, Kukutla P, *et al.* (2011) Dynamic gut microbiome across life history of the malaria mosquito *Anopheles gambiae* in Kenya. *PLoS ONE*, **6**, e24767.

Wharton RH (1962) The biology of Mansonia mosquitoes in relation to the transmission of filariasis in Malaya. *Bulletin of the Institute of Medical Research, Malaya*, **11**, 1–114.

JOP and FMJ designed the experiment; JOP performed the experiment; CMM provided field advice and logistics; JOP and WJP involved in bioinformatic analyses; JOP and FMJ analysed data and wrote paper.

Data accessibility

Raw reads from 454 GS FLX: NCBI SRA Accession SRR516961.

Filtered data used in manuscript (annotated bacterial OTUs, sample locations): Dryad Digital Repository doi: 10.5061/dryad.2dm26.

Supporting information

Additional Supporting Information may be found in the online version of this article.

Table S1 Multiplier Identifier (MID) sequences.

Table S2 The classification of OTUs (Fig. 2A) showing closest family and/or genus.

Table S3 Distribution of all classified bacteria between mosquito host species. Bacteria are classified to genus level with RDP II Classifier (Wang *et al.* 2007).

Fig. S1 Map of Kenya showing Kilifi and Malindi: the two districts from which mosquitoes were collected. Jaribuni and Mbogolo are inland villages in both districts respectively while the other sampling sites nearer to the coast.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.